Chapter 1. Labor Force Data Derived from the Current Population Survey

Each month, the Bureau analyzes and publishes statistics on the labor force, employment, and unemployment, classified by a variety of demographic, social, and economic characteristics. These statistics are derived from the Current Population Survey (CPS), which is conducted by the Bureau of the Census for BLS. This monthly survey of the population is conducted using a scientifically selected sample of households, representative of the civilian noninstitutional population of the United States.

IN THIS CHAPTER

Background

Specific concepts of the labor force, employment, and unemployment were developed in the later stages of the depression of the 1930s. Before the 1930s, aside from attempts in some of the decennial censuses, no direct measurements were made of the number of jobless persons. Mass unemployment in the early 1930s increased the need for statistics, and widely conflicting estimates based on a variety of indirect techniques began to appear. Dissatisfied with these methods, many research groups, as well as State and municipal governments, began experimenting with direct surveys or samples of the population. In these surveys, an attempt was made to classify the population as employed, unemployed, or out of the labor force by means of a series of questions addressed to each individual. In most of the surveys, the employed were defined as persons with occupations ("gainful workers"), and the unemployed were defined as those who were not working but were "willing and able to work." These concepts did not meet the standards of objectivity that many technicians felt were necessary to measure either the level of unemployment at a point in time or changes over periods of time. Counts of gainful workers did not have a current dimension, and the criterion "willing and able to work," when applied in specific situations, appeared to be too intangible and too dependent upon the interpretation and attitude of the persons being interviewed.

A set of precise concepts was developed in the late 1930s to meet these various criticisms. The classification of an individual depended principally upon his or her actual *activity* within a designated time period, that is, was it the individual working, looking for work, or engaged in other activities? These concepts were adopted for the national sample survey of households, called the Monthly Report of Unemploy-ment, initiated in 1940 by the Works Progress Administration.

The household survey was transferred to the Bureau of the Census in late 1942, and its name was changed to the Monthly Report on the Labor Force. The survey title was changed once more in 1948 to the present Current Population Survey in order to reflect its expanding role as a source for a wide variety of demographic, social, and economic characteristics of the population. In 1959, responsibility for analyzing and publishing the CPS labor force data was transferred to BLS; the Bureau of the Census continues to collect the data.

Description of Survey

The CPS provides statistics on the labor force status of the civilian noninstitutional population 16 years of age and over. Persons under 16 years of age are excluded from the official definition of the labor force because child labor laws, compulsory school attendance, and general social custom prevent most of these children from working in the United States. The institutional population, which also is excluded from coverage, consists of inmates of penal and mental institutions, sanitariums, and homes for the aged, infirm, and needy.

The CPS is collected each month from a probability sample of approximately 50,000 occupied households. Respondents are assured that all information obtained is completely confidential and is used only for the purpose of statistical analysis. Although the survey is conducted on a strictly voluntary basis, refusals to cooperate amount to about 4 percent each month.

The time period covered in the monthly survey is a calendar week. A calendar week was selected as the survey reference period because the period used must be short enough so that the data obtained are "current" but not so short that the occurrence of holidays or other accidental events might cause erratic fluctuations in the information obtained. A calendar week fulfills these conditions as well as being a convenient and easily defined period of time. Since July 1955, the calendar week, Sunday through Saturday, which includes the 12th day of the month has been defined as the reference week. The actual survey is conducted during the following week, which is the week containing the 19th day of the month.

Concepts

The criteria used in classifying persons on the basis of their labor force activity and some of the major statistics obtained from the CPS are as follows:

Employed persons. All those who, during the reference week, (1) did any work at all as paid employees, worked in their own business, profession, or on their own farm, or who worked 15 hours or more as unpaid workers in a family-operated enterprise; and (2) all those who did not work but had jobs or businesses from which they were temporarily absent due to illness, bad weather, vacation, child-care problems, labor dispute, maternity or paternity leave, or other family or personal obligations—whether or not they were paid by their employers for the time off and whether or not they were seeking other jobs. Each employed person is counted only once, even if they hold more than one job. Multiple jobholders are counted in the job at which they worked the greatest number of hours during the reference week. Included in the total are employed citizens of foreign countries who are residing in the United States, but are not living on the premises of an embassy. Excluded are persons whose only activity consisted of work around their own home (such as housework, painting, repairing, etc.) or volunteer work for religious, charitable, and similar organizations.

Unemployed persons. All persons who had no employment during the reference week, were available for work, except for temporary illness, and had made specific efforts, such as contacting employers, to find employment sometime during the 4-week period ending with the reference week. Persons who were waiting to be recalled to a job from which they had been laid off need not have been looking for work to be classified as unemployed.

Duration of unemployment represents the length of time (through the current reference week) that persons classified as unemployed had been continuously looking for work and thus is a measure of an in-progress spell of joblessness. For persons on layoff, duration of unemployment represents the number of full weeks since the end of their most recent period of employment. Two useful measures of the duration of unemployment are the mean and the median. Mean duration is the arithmetic average computed from single weeks of unemployment. Median duration is the midpoint of a distribution of weeks of unemployment.

The reasons for unemployment are divided into five major groups. (1) Job losers, comprised of (a) persons on temporary layoff, who have been given a date to return to work or who expect to return within 6 months (persons on layoff need not be looking for work to be classified as unemployed), and (b) permanent job losers, whose employment ended involuntarily and who began looking for work; (2) Job leavers, persons who quit or otherwise terminated their employment voluntarily and immediately began looking for work; (3) Persons who completed temporary jobs, who began looking for work after the jobs ended; (4) Reentrants, persons who previously worked but were out of the labor force prior to beginning their job search; and (5) New entrants, persons who never worked but were searching for work.

Civilian labor force. The civilian labor force comprises the total of all civilians classified as employed and unemployed.

Unemployment rate. The unemployment rate represents the proportion of the civilian labor force that is unemployed.

Participation rate. This represents the proportion of the population that is in the labor force.

Employment-population ratio. This represents the proportion of the population that is employed.

Not in the labor force. Included in this group are all persons in the civilian noninstitutional population who are neither employed nor unemployed. Information is collected on their desire for and availability to take a job at the time of the CPS interview, job search activity in the prior year,

and reason for not looking in the 4-week period ending with the reference week. This group includes persons marginally attached to the labor force, defined as persons not in the labor force who want and are available for a job and who have looked for work sometime in the past 12 months (or since the end of their last job if they held one within the past 12 months), but are not currently looking. Those marginally attached to the labor force are divided into those not currently looking because they believe their search would be futile and those not currently looking for other reasons. For discouraged workers, the main reason for not recently looking for work was one of the following: Believes no work available in line of work or area; couldn't find any work; lacks necessary schooling, training, skills or experience; employers think too young or too old; or other types of discrimination.

Multiple jobholders. These are employed persons who, during the reference week, had either two or more jobs as a wage and salary worker, or were self-employed and also held a wage and salary job, or worked as an unpaid family worker and also held a wage and salary job.

At work part time for economic reasons. Sometimes referred to as involuntary part time, this category refers to individuals who gave an economic reason for working 1 to 34 hours during the reference week. Economic reasons include slack work or unfavorable business conditions, inability to find full-time work, and seasonal declines in demand. Those who usually work part time must also indicate that they want and are available to work full time to be classified as on part time for economic reasons.

At work part time for noneconomic reasons. This group includes those persons who usually work part time and were at work 1 to 34 hours during the reference week for a noneconomic reason. Noneconomic reasons include, for example: Illness or other medical limitations, child-care problems or other family or personal obligations, school or training, retirement or Social Security limits on earnings, and being in a job where full-time work is less than 35 hours. This group also includes those who gave an economic reason for usually working 1 to 34 hours but said they do not want to work full time or were unavailable for such work.

Usual full- or part-time status. Full-time workers are those who usually worked 35 hours or more (at all jobs combined). This group will include some individuals who worked less than 35 hours in the reference week for either economic or noneconomic reasons and those who are temporarily absent from work. Part-time workers are those who usually work less than 35 hours per week (at all jobs), regardless of the number of hours worked in the reference week. This may include some individuals who actually worked more than 34 hours in the reference week, as well as those temporarily absent from work who usually work less than 35 hours.

Usual weekly earnings for wage and salary workers. Data are collected on earnings before taxes and other deductions, and include any overtime pay, commissions, or tips usually received (at the main job in the case of multiple jobholders). Earnings reported on a basis other than weekly (e.g., annual, monthly, hourly) are converted to weekly. The term "usual" is as perceived by the respondent. If the respondent asks for a definition of usual, interviewers are instructed to define the term as more than half the weeks worked during the past 4 or 5 months.

Redesign of the Current Population Survey

A major redesign of the CPS was implemented in January 1994, with the primary objective to improve the quality of the data derived from the survey by introducing a new questionnaire and modernized data collection methods. Prior to the redesign, the survey questionnaire had been virtually unchanged since 1967, at which time changes had been introduced based on recommendations of the Gordon Committee (President's Committee to Appraise Employment and Unemployment, 1962). Additional changes were proposed in the late 1970's based on the Levitan Commission (National Commission on Employment and Unemployment, 1979); these, in part, formed the basis for the redesign of the questionnaire.

The redesign of the questionnaire had four main objectives: 1) To adopt a computer-assisted interviewing environment, 2) to measure the official labor force concepts more precisely, 3) to expand the amount of data available, and 4) to implement several definitional changes.

Computerization

The new questionnaire was designed for a computer-assisted interview. In most cases, interviewers conduct the survey either in person at the respondent's home or by telephone from the interviewer's home, using laptop computers on which the questionnaire has been programmed. This mode of data collection is known as computer-assisted personal interviewing (CAPI). Interviewers ask the survey questions as they appear automatically on the screen of the laptop, and then type the responses directly into the computer.

Computer-assisted interviewing has important benefits, most notably: automation facilitates the use of a relatively complex questionnaire, incorporating complicated skip patterns and standardized followup questions. Yet, certain questions are automatically tailored to the individual's situation to make them more understandable to the respondent. The computerized questionnaire also has several built-in editing features, including automatic checks for internal consistency

and unlikely responses. An automated interview also permits dependent interviewing, that is, the use of information in the current interview that was obtained in a previous month's interview. Dependent interviewing reduces respondent and interviewer burden, while improving consistency of the data from one month to the next. The technique is being used to confirm the previously reported occupation of a person's job and, for many people not in the labor force, to confirm their status as retired or disabled.

Major questionnaire changes

While the labor force status of most people is straightforward, some persons are more difficult to classify correctly, especially if they are engaged in activities that are relatively informal or intermittent. Many of the changes to the questionnaire were made to deal with such cases. This was accomplished by rewording and adding questions to conform more precisely with the official definitions, making the questions easier to understand and answer, minimizing reliance on volunteered responses, revising response categories, and taking advantage of the benefits of an automated interview. Areas affected by these improvements include:

On layoff. Persons on layoff are defined as those who are separated from a job to which they are awaiting recall. The old questionnaire, however, was not structured to consistently obtain information on the expectation of recall. In order to measure layoffs more accurately, questions were added to determine if people reported to be on layoff did in fact have an expectation of recall—that is, had they been given a specific date to return to work or, at least, had they been given an indication that they would be recalled within the next 6 months.

Jobsearch methods. To allow interviewers to better distinguish between active and passive methods, the response categories for jobsearch methods were expanded and reformatted. Also, the basic question on jobsearch methods was reworded and followup questions were added to encourage respondents to report all types of jobsearch activity.

Hours at work. To improve the accuracy of these data, the series of questions on hours worked was reordered to incorporate a recall strategy, which asks for usual hours first, then about possible time taken off or extra hours worked during the reference week, and finally about hours actually worked.

Reasons for working part time. Persons who work part time do so either for noneconomic reasons (that is, because of personal constraints or preferences) or for economic reasons (that is, because of business-related reasons such as slack work or the lack of full-time opportunities). Because respondents typically are not familiar with this distinction, the question was reworded to provide examples of the two types of reasons. More importantly, the measurement of working part time involuntarily (or for economic reasons) was modified to better reflect the concept. Starting in 1994, workers who usually work part time and are working part time involuntarily must want and be available for full-time work.

Earnings. With the former questionnaire, respondents were asked to report their earnings as a weekly amount, even though that may not have been the easiest way to recall or report their earnings. In the new version, respondents are asked to report earnings in the time frame which they find easiest, for example, hourly, weekly, biweekly, monthly, or annual. Weekly earnings are automatically calculated for persons who respond on a basis other than weekly.

New data and definitional changes

The questionnaire redesign also made it possible to collect several types of data regularly for the first time, namely:

Multiple jobholding. Employed persons are now asked each month whether they had more than one job. This allows BLS to produce estimates of multiple jobholding on a monthly basis, rather than having to derive them through special, periodic supplements.

Usual hours. All employed persons are asked each month about the hours they usually work. Previously, information on usual hours was collected from just one-quarter of wage and salary workers each month.

In addition, several labor force definitions were modified. The most important definitional changes concerned discouraged workers. The Levitan Commission had criticized the former definition, because it was based on a subjective desire for work and on somewhat arbitrary assumptions about an individual's availability to take a job. As a result of the redesign, two requirements were added: For persons to qualify as discouraged, they must have engaged in some jobsearch within the past year (or since they last worked if they worked within the past year), and they must be currently available to take a job. (Formerly, availability was inferred from responses to other questions; now there is a direct question.) Also, beginning in January 1994, questions on this subject are asked of the full CPS sample, permitting estimates of the number of discouraged workers to be published monthly (rather than quarterly).

Another important definitional change concerned unemployed persons who were not working just before their jobsearch commenced, that is, new entrants or reentrants to the labor force. Prior to 1994, new entrants were defined as jobseekers who had never worked at a full-time job lasting 2 weeks or longer; reentrants were defined as jobseekers who had held a full-time job for at least 2 weeks and had then spent some time out of the labor force prior to their most recent period of jobsearch. These definitions have been modified to encompass any type of job, not just a full-time job of at least 2 weeks duration. Thus, new entrants are now

defined as jobseekers who have never worked at all, and reentrants are jobseekers who have worked before, but not immediately prior to their jobsearch.

Introduction of 1990 Census-based population controls

Beginning with the CPS estimates for January 1994, the independent national population controls used for the age-sex-race groups in the second-stage estimation procedure were prepared by projecting forward the resident population as enumerated on April 1, 1990. (See section on Sampling.) Also, estimates of the decennial census undercount, obtained from the Post Enumeration Survey (PES), were added to the population controls. In 1996, CPS estimates for 1990-93 were revised to incorporate these 1990 census-based population controls.

Effect of new population controls. The effect of the new population controls on the national estimates for 1990-93 were very consistent with the effects described in the article, "Revisions in the Current Population survey Effective January 1994," appearing in the February 1994 issue of Employment and Earnings. Increases in levels were concentrated in younger age groups, races other than white or black, and persons of Hispanic origin. Data for older age groups generally decreased as a result of these revisions.

Over the 4 years subject to revision, the annual average estimates of the civilian noninstitutional population 16 years and over increased at a gradual rate, beginning with an increase of about 1.1 million (0.6 percent) in 1990, and ending with an increase of about 1.3 million (0.7 percent) in 1993. Because of the demographic distribution of the population increase, particularly with respect to age, the estimated levels of civilian labor force rose by almost as much as the population, resulting in average increases of 0.1 in overall participation rates and employment-population ratios. Within the civilian labor force, the new controls caused average increases of 0.1 in the overall unemployment rate because demographic groups with relatively high unemployment rates, such as youth and persons of Hispanic origin, had relatively large upward population adjustments.

Sampling

Since the inception of the survey, there have been various changes in the design of the CPS sample. The sample is traditionally redesigned and a new sample selected after each decennial census. Also, the number of sample areas and the number of sample persons are changed occasionally. Most of these changes are made in order to improve the efficiency of the sample design, increase the reliability of the sample estimates, or control cost. Since the mid-1980's, the CPS has had a State-based sample design, meaning all sampling operations such as allocation and selection were implemented at the State level.

A redesigned CPS sample based on the 1990 decennial census was selected for use during the 1990s. Households from this new sample were phased into the CPS between April 1994 and July 1995. The July 1995 sample was the first monthly sample based entirely on the 1990 census.

The original 1990 census-based sample design included about 66,000 housing units per month located in 792 selected geographic areas called primary sampling units (PSU's). The sample was initially selected to meet specific reliability criteria for the Nation, for each of the 50 States and the District of Columbia, and for the sub-State areas of New York City and the Los Angeles-Long Beach metropolitan area. In 1996, the original sample design reliability criteria were modified to reduce costs. The current criteria, given below, are based on the coefficient of variation (CV) of the unemployment level, where the CV is defined as the standard error of the estimate divided by the estimate, expressed as a percentage. These CV controls assume a 6-percent unemployment rate in order to establish a consistent specification of sampling error.

The current sample design, introduced in January 1996, includes about 59,000 households from 754 sample areas and maintains a 1.9-percent CV on national monthly estimates of unemployment level. This translates into a change of 0.2 percentage point in the unemployment rate being significant at a 90-percent confidence level. For each of the 50 States and for the District of Columbia, the design maintains a CV of at most 8-percent on the annual average estimate of unemployment level, assuming a 6-percent unemployment rate. Due to the national reliability criterion, estimates for several large States are substantially more reliable than the State design criterion requires. Annual average unemployment estimates for California, Florida, New York, and Texas, for example, carry a CV of than 4 percent.

In the first stage of sampling, the 754 sample areas (PSU's) are chosen. In the second stage, ultimate sampling unit clusters composed of about four housing units each are selected. Each month, about 59,000 housing units are assigned for data collection, of which about 50,000 are occupied and thus eligible for interview. The remainder are units found to be destroyed, vacant, converted to nonresidential use, containing persons whose usual place of residence is elsewhere, or ineligible for other reasons. Of the 50,000 housing units, about 6 to 7 percent are not interviewed in a given month due to temporary absence (vacation, etc.), other failures to make contact after repeated attempts, inability of persons contacted to respond, unavailability for other reasons, and refusals to cooperate (about half of the noninterviews). Information is obtained each month for about 94,000 persons 16 years of age or older.

Selection of sample areas. The entire area of the United States, consisting of 3,141 counties and independent cities, is divided into 2,007 sample units (PSU's). PSU's are defined within States and do not cross State boundaries. In

most States, a PSU consists of a county or a number of contiguous counties. In New England and Hawaii, minor civil divisions are used instead of counties.

Metropolitan areas within a State are used as a basis for forming many PSU's. Outside of metropolitan areas, two or more counties normally are combined to form a PSU except when the geographic area of an individual county is too large. Combining counties to form PSU's provides greater heterogeneity; a typical PSU includes urban and rural residents of both high and low economic levels and encompasses, to the extent feasible, diverse occupations and industries. Another important consideration is that the PSU be sufficiently compact so that, with a small sample spread throughout, it can be efficiently canvassed without undue travel cost.

The 2,007 PSU's are grouped into strata within each State. Then one PSU is selected from each stratum with the probability of selection proportional to the population of the PSU. Nationally, there are a total of 428 PSU's in strata by themselves. These strata are self-representing and are generally the most populous PSU's in each State. The 326 remaining strata are formed by combining PSU's that are similar in such characteristics as unemployment, proportion of housing units with three or more persons, number of persons employed in various industries, and average monthly wages for various industries. The single PSU randomly selected from each of these strata is non-self-representing because it represents not only itself but the entire stratum. The probability of selecting a particular PSU in a non-self-representing stratum is proportional to its 1990 population. For example, within a stratum, the chance that a PSU with a population of 50,000 would be selected for the sample is twice that for a PSU having a population of 25,000.

Selection of sample households. Because the sample design is State based, the sampling ratio differs by State and depends on State population size as well as both national and State reliability requirements. The State sampling ratios range roughly from 1 in every 100 households to 1 in every 3,000 households. The sampling ratio occasionally is modified slightly to hold the size of the sample relatively constant given the overall growth of the population (called "sample maintenance reduction"). The sampling ratio used within a sample PSU depends on the probability of selection of the PSU and the sampling ratio for the State. In a sample PSU with a probability of selection of 1 in 10 and a State sampling ratio of 3,000, a within-PSU sampling ratio of 1 in 300 achieves the desired overall ratio of 1 in 3,000 for the stratum.

The 1990 within-PSU sample design was developed using block-level data from the 1990 census. (The 1990 census was the first decennial census that produced data at the block level for the entire country.) Normally, census blocks are bounded by streets and other prominent physical features such as rivers or railroad tracks. County, Minor Civil Division, and census place limits also serve as block boundaries. In cities, blocks can be bounded by four streets and be quite small in land area. In rural areas, blocks can be several square miles in size.

For the purpose of sample selection, census blocks were grouped into three strata: Unit, group quarters, and area. (Occasionally, units within a block were split between the unit and group quarters strata.) The unit stratum contained regular housing units with addresses that were easy to locate (e.g., most single family homes, townhouses, condominiums, apartment units, and mobile homes). The group quarters stratum contained housing units where residents shared common facilities or received formal or authorized care or custody. Unit and group quarters blocks exist primarily in urban and suburban areas. The area stratum contains blocks with addresses that are more difficult to locate. Area blocks exist primarily in rural areas.

To reduce the variability of the survey estimates and to ensure that the within-PSU sample would reflect the demographic and socioeconomic characteristics of the PSU, blocks within the unit, group quarters, and area strata were sorted using geographic and block-level data from the census. Examples of the census variables used for sorting include proportion of minority renter-occupied housing units, proportion of housing units with female householders, and proportion of owner-occupied housing units. The specific sorting variables used differed by type of PSU (urban or rural) and stratum.

Within each block, housing units were sorted geographically and grouped into clusters of approximately four units. A systematic sample of these clusters was then selected independently from each stratum using the appropriate within-PSU sampling ratio. The geographic clustering of the sample units reduces field representative travel costs. Prior to interviewing, special listing procedures are used to locate the particular sample addresses in the group quarters and area blocks.

Units in the three strata described above all existed at the time of the 1990 decennial census. Through a series of additional procedures, a sample of building permits is included in the CPS to represent housing units built after the decennial census. Adding these newly built units keeps the sample up-to-date and representative of the population. It also helps to keep the sample size stable: over the life of the sample, the addition of newly built housing units compensates for the loss of "old" units which may be abandoned, demolished, or converted to nonresidential use. In normal circumstances, the number of eligible households in the sample grows slowly. Sample maintenance reduction procedures are periodically implemented to hold the size of the sample relatively constant.

Rotation of sample. Part of the sample is changed each month. Each monthly sample is divided into eight representative subsamples or rotation groups. A given rotation group is interviewed for a total of 8 months, divided into

two equal periods. It is in the sample for 4 consecutive months, leaves the sample during the following 8 months, and then returns for another 4 consecutive months. In each monthly sample, one of the eight rotation groups is in the first month of enumeration, another rotation group is in the second month, and so on. (The rotation group in the fifth month of enumeration is returning after an 8-month break.) Under this system, 75 percent of the sample is common from month to month and 50 percent from year to year for the same month. This procedure provides a substantial amount of month-to-month and year-to-year overlap in the sample, thus providing better estimates of change and reducing discontinuities in the series of data without burdening any specific group of households with an unduly long period of inquiry.

Collection Methods

Each month, during the calendar week containing the 19th day, interviewers contact a "responsible" person in each of the sample households in the CPS. At the time of the first enumeration of a household, the interviewer visits the household and prepares a roster of the household members, including their personal characteristics (date of birth, sex, race, ethnic origin, marital status, educational attainment, veteran status, etc.) and their relationship to the person maintaining the household. This information is entered by the interviewers into laptop computers. This roster is then checked for accuracy and brought up to date at each subsequent interview to take account of new or departed residents, changes in marital status, and similar items. The information on personal characteristics is thus available each month for identification purposes and for cross-classification with economic characteristics of the sample population.

Each month, one-fourth of the households in the sample are changed, so that no household is interviewed more than 4 consecutive months. This practice avoids placing too heavy a burden on the households selected for the sample. After a household is interviewed for 4 consecutive months, it leaves the sample for 8 months and then is again interviewed for the same 4 calendar months a year later, before leaving the sample for good.

Personal visits are required in the first and fifth months that the household is in the sample. In other months, the interview may be conducted by telephone if the respondent agrees to this procedure. Also, if no one is at home when the interviewer visits, the respondent may be contacted by telephone after the first month. Approximately 60 percent of the households in any given month are interviewed by telephone. A portion of the households (12 percent) is interviewed via computer-assisted telephone interviewing (CATI), from three centralized telephone centers (located in Hagerstown, Maryland; Jefferson, Indiana; and Tucson, Arizona) by interviewers who also use a computerized questionnaire.

At each monthly visit, each household member 16 years of age and over is asked a series of standard questions on economic activity during the preceding week. The primary purpose of these questions is to classify the sample population into the three basic economic groups: The employed, the unemployed, and those not in the labor force.

Additional questions are asked each month to help clarify the information on labor force status. For the employed, information is obtained on hours worked during the survey week, a description of their current job, and whether they work at more than one job. For those temporarily away from their jobs, the enumerator records their reason for not working during the survey week, whether or not they were paid for their time off, and whether they usually work full or part time. For the unemployed, information is obtained on (1) the method(s) used to find work during the 4 weeks prior to the interview, (2) the reasons the unemployed persons had started to look for work, (3) the length of time they had been looking for work, (4) whether they were seeking full-or part-time work, and (5) a description of their last full-time civilian job. For persons not in the labor force, questions are asked each month about their desire for and availability to take a job, job search activity in the prior year, and reason for not looking in the 4 week period prior to the survey week. In addition, all employed persons in households in the outgoing rotation groups are asked questions on union status and hourly and weekly earnings; questions also are asked on the work history and jobseeking intentions of individuals not in the labor force.

At the end of each day's interviewing, the data collected are transmitted to the Census Bureau's central computer in Washington, DC. Once files are transmitted to the main computer, they are deleted from the laptops.

Because of the crucial role interviewers have in the household survey, a great amount of time and effort is spent maintaining the quality of their work. Interviewers are given intensive training, including classroom lectures, discussion, practice, observation, home-study materials, and on-the-job training. At least once a year, they convene for day-long training and review sessions, and, also at least once a year, they are accompanied by a supervisor during a full day of interviewing to determine how well they carry out their assignments.

A selected number of households are reinterviewed each month to determine whether the information obtained in the first interview was correct. The information gained from these interviews is used to improve the entire training program.

Estimation Methods

Under the estimating methods used in the CPS, all of the results for a given month become available simultaneously and are based on returns from the entire panel of respondents. The estimation procedure involves weighting the

data from each sample person by the inverse of the probability of the person being in the sample. This gives a rough measure of the number of actual persons that the sample person represents. Since 1985, most sample persons within the same State have had the same probability of selection. Some selection probabilities may differ within a State due to the sample design or for operational reasons. Field subsampling, for example, which is carried out when areas selected for the sample are found to contain many more households than expected, may cause probabilities of selection to differ for some sample areas within a State. Through a series of estimation steps (outlined below), the selection probabilities are adjusted for noninterviews and survey undercoverage; data from previous months are incorporated into the estimates through the composite estimation procedure.

1. *Noninterview adjustment.* The weights for all interviewed households are adjusted to account for occupied sample households for which no information was obtained because of absence, impassable roads, refusals, or unavailability of the respondents for other reasons. This noninterview adjustment is made separately for clusters of similar sample areas that are usually, but not necessarily, contained within a State. Similarity of sample areas is based on Metropolitan Statistical Area (MSA) status and size. Within each cluster, there is a further breakdown by residence. Each MSA cluster is split by "central city" and "balance of the MSA." Each non-MSA cluster is split by "urban" and "rural" residence categories. The proportion of sample households not interviewed varies from 6 to 7 percent, depending on weather, vacation, etc.

2. *Ratio estimates*. The distribution of the population selected for the sample differs somewhat, by chance, from that of the population as a whole in such characteristics as age, race, sex, and State of residence. Because these characteristics are closely correlated with labor force participation and other principal measurements made from the sample, the survey estimates can be substantially improved when weighted appropriately by the known distribution of these population characteristics. This is accomplished through two stages of ratio adjustment, as follows:

   a. *First-stage ratio estimation*. The purpose of the first-stage ratio adjustment is to reduce the contribution to variance that results from selecting a sample of PSU's rather than drawing sample households from every PSU in the Nation. This adjustment is made to the CPS weights in two race cells: Black and nonblack; it is applied only to data from PSU's that are not self-representing and for those States that have a substantial number of black households. The procedure corrects for differences that existed in each State cell at the time of the 1990 census between 1) the race distribution of the population in sample PSU's and 2) the race distribution of all PSU's (both 1 and 2 exclude self-representing PSU's).

   b. *Second-stage ratio estimation*. This procedure substantially reduces the variability of estimates and corrects, to some extent, for CPS undercoverage. The CPS sample weights are adjusted to ensure that sample-based estimates of population match independent population controls. Three sets of controls are used:

   1) 51 State controls for the civilian noninstitutional population 16 years of age and older.

   2) National civilian noninstitutional population con trols for 14 Hispanic and 5 non-Hispanic age-sex categories.

   3) National civilian noninstitutional population conrols for 66 white, 42 black, and 10 "other" categories.

   The independent population controls are prepared by projecting forward the resident population as enumerated on April 1, 1990. The projections are derived by updating demographic census data with information from a variety of other data sources that account for births, deaths, and net migration. Subtracting estimated numbers of resident Armed Forces personnel and institutionalized persons reduces the resident population to the civilian noninstitutional population. Estimates of net census undercount, determined from the Post Enumeration Survey, are added to the population projections. Prior to January 1994, the projections were based on earlier censuses, and there was no correction for census undercount. A summary of the current procedures used to make population projections is given in "Revisions in the Current Population Survey Effective January 1994," appearing in the February 1994 issue of *Employment and Earnings.*

3. *Composite estimation procedure*. The last step in the preparation of most CPS estimates makes use of a composite estimation procedure. The composite estimate is based on a weighted average of two factors: (1) The two-stage ratio estimate based data from the entire sample for the current month; and (2), the composite estimate for the previous month, adjusted by an estimate of the month-to-month change based on the six rotation groups common to both months. A bias adjustment term is added to the weighted average to reduce variance and partially account for bias associated with month-in-sample estimates. This month-in-sample bias is exhibited by unemployment estimates for persons in their first and fifth months in the CPS being generally higher than estimates obtained for the other months.

   Composite estimation results in a reduction in the sampling error beyond that which is achieved through the two

stages of ratio estimation. For some items, the reduction is substantial. The resultant gains in reliability are greatest in estimates of month-to-month change, although gains are also usually obtained for estimates of level in a given month, change from year to year, and change over other intervals of time.

Seasonal Adjustment

Over the course of a year, the size of the Nation's labor force, the levels of employment and unemployment, and other measures of labor market activity undergo sharp fluctuations due to such seasonal events as changes in weather, reduced or expanded production, harvests, major holidays, and the opening and closing of schools. Because these seasonal events follow a more or less regular pattern each year, their influence on statistical trends can be eliminated by adjusting the statistics from month to month. These adjustments make it easier to observe the cyclical and other nonseasonal movements in the series. In evaluating changes in a seasonally adjusted series, it is important to note that seasonal adjustment is merely an approximation based on past experience. Seasonally adjusted estimates have a broader margin of possible error than the original data on which they are based, because they are subject not only to sampling and other errors but are also affected by the uncertainties of the seasonal adjustment process itself.

Since January 1980, national labor force data have been seasonally adjusted with a procedure called *X-11 ARIMA* (Auto-Regressive Integrated Moving Average), which was developed at Statistics Canada as an extension of the standard X-11 method.

At the beginning of each calendar year, projected seasonal adjustment factors are calculated for use during the January-June period. In July of each year, BLS calculates and publishes in *Employment and Earnings* projected seasonal adjustment factors for use in the second half, based on the experience through June. Revisions of historical data, usually for the most recent 5 years, are made only at the beginning of each calendar year. However, as a result of revisions to the estimates for 1970-81 based on 1980 census population counts, revisions to seasonally adjusted series in early 1982 were carried back to 1970. In 1994, data were revised only for that year because of the major redesign and 1990 census-based population controls, adjusted for the estimated undercount, introduced into the Current Population Survey. In 1996, 1990-93 data also were revised to incorporate these 1990 census-based population controls and seasonally adjusted series were revised back to 1990.

Presentation and Uses

The CPS provides comprehensive information on the social, demographic, and economic characteristics of the civilian noninstitutional population 16 years of age and over, the employed, and the unemployed.

Each month, the employment and unemployment data are published initially in The Employment Situation news release about 2 weeks after they are collected. The release includes a narrative summary and analysis of the major employment and unemployment developments together with tables containing statistics for the principal data series. The news release is available on the Internet and can be accessed via the World Wide Web. The Universal Resource Locator is: http://stats.bls.gov/cpshome.htm. The news release also is available on the BLS fax-on-demand service.

More detailed statistics are subsequently published in Employment and Earnings, a monthly periodical. The detailed tables provide information on the labor force, employment, and unemployment by a number of characteristics, such as age, sex, race, marital status, industry, and occupation. In addition, the January issue of Employment and Earnings provides annual averages on employment and earnings by detailed occupational categories, union affiliation, and employee absences.

About 32,000 of the monthly labor force data series plus quarterly and annual averages are maintained in LABSTAT, the Bureau's public database, on the Internet. They can be accessed from http:/stats.bls.gov/cgi-bin/dsrv?lf. In most cases, these data are available from the inception of the series through the current month.

The CPS is used also for a program of special inquiries to obtain detailed information from particular segments, or for particular characteristics of the population and labor force. About four such special surveys are made each year. The inquiries are repeated annually in the same month for some topics, including the earnings and total incomes of individuals and families (published by the Bureau of the Census); the extent of work experience of the population during the calendar year; and the employment of school-age youth, high school graduates and dropouts, and recent college graduates. Surveys also are made periodically on subjects such as contingent work, job mobility, job tenure, displaced workers, and disabled veterans.

Generally, the persons who provide information for the monthly CPS questions also answer the supplemental questions. Occasionally, the kind of information sought in the special survey requires the respondent to be the person about whom the questions are asked. The results of these special surveys are usually first published as news releases and subsequently in the *Monthly Labor Review* or BLS reports.

In addition to the regularly tabulated statistics described above, special data can be generated through the use of the CPS individual record (micro) files. These files contain records of the responses to the survey questionnaire for all individuals in the survey. While the files can be used simply to create additional cross-sectional detail, an important feature of their use is the ability to match the

records of specific individuals at different points in time during their participation in the survey. By matching these records, data files can be created which lend themselves to some limited longitudinal analysis and the investigation of short-run labor market dynamics. An example is the statistics on gross labor force flows, which indicate how many persons move among the labor force status categories each month. Although a number of technical difficulties lie in the path of more complete use of these data files for the purposes of longitudinal analysis, this area is continually being investigated and holds considerable promise. Microdata files are available for all months since January 1976 and for various months in prior years. These data are made available on magnetic tape, CD-Rom, or diskette. Address inquires regarding these files to: Division of Data Development, Bureau of Labor Statistics, Room 4965, 2 Massachusetts Ave., NE, Washington, DC 20212-0001, phone 202-606-6345.

Limitations

Geographic. Although the present CPS sample is a State-based design, the CPS produces reliable national monthly estimates. The sample does not permit the production of reliable monthly estimates for the States. However, demographic, social, and economic detail are published annually for the census regions and divisions, all States and the District of Columbia, 50 large metropolitan areas, and selected central cities. The production of subnational labor force and unemployment estimates is discussed in more detail in chapter 4 of this bulletin.

Sources of errors in the survey estimates. There are two types of errors possible in an estimate based on a sample survey—sampling and nonsampling. The mathematical discipline of sampling theory provides methods for estimating standard errors when the probability of selection of each member of a population can be specified. The standard error, a measure of sampling variability, can be used to compute confidence intervals that indicate a range of differences from true population values that can be anticipated because only a sample of the population has been surveyed. Nonsampling errors such as response variability, response bias, and other types of bias occur in complete censuses as well as sample surveys. In some instances, nonsampling error may be more tightly controlled in a well-conducted survey where it is feasible to collect and process the data more skillfully. Reinterview programs are often used to measure response variability and response bias. Estimation of other types of bias is one of the most difficult aspects of survey work, and often adequate measures of bias cannot be made.

Nonsampling error. The full extent of nonsampling error is unknown, but special studies have been conducted to quantify some sources of nonsampling error in the CPS. The effect of nonsampling error should be small on estimates of relative change, such as month-to-month change. Estimates of monthly levels would be more severely affected by nonsampling error.

Nonsampling errors in surveys can be attributed to many sources, e.g., the inability to obtain information about all persons in the sample; differences in the interpretation of questions; inability or unwillingness of respondents to provide correct information; inability to recall information; errors made in collecting and processing the data; errors made in estimating values for missing data; and failure to represent all sample households and all persons within sample households (undercoverage).

Nonsampling errors occurring in the interview phase of the survey are studied by means of a reinterview program. This program is used to estimate various sources of error as well as to evaluate and control the work of the interviewers. A random sample of each interviewer's work is inspected through reinterview at regular intervals. The results indicate, among other things, that the data published from the CPS are subject to moderate systematic biases.

The effects of some components of nonsampling error in the CPS data can be examined as a result of the rotation plan used for the sample, since the level of the estimates varies by rotation group. For example, unemployment estimates from a rotation group tend to be higher in the first and fifth months of interviewing.

Undercoverage in the CPS results from missed housing units and missed persons within sample households. The CPS covers about 92 percent of the decennial census population (adjusted for the undercount). It is known that the CPS undercoverage varies with age, sex, race, and Hispanic origin. Generally, undercoverage is larger for men than for women and larger for blacks, Hispanics, and other races than for whites. Ratio adjustment to independent age-sex-race-origin population controls, as described previously, partially corrects for the biases due to survey undercoverage. However, biases exist in the estimates to the extent that missed persons in missed households or missed persons in interviewed households have different characteristics than interviewed persons in the same age-sex-race-origin group.

The independent population estimates used in the estimation procedure may be a source of error although, on balance, their use substantially improves the statistical reliability of many of the figures. Errors may arise in the independent population estimates because of underenumeration of certain population groups or errors in age reporting in the last census (which serves as the base for the estimates) or similar problems in the components of population change (mortality, immigration, etc.) since that date.

Sampling error. When a sample rather than the entire population is surveyed, estimates differ from the true population values that they represent. This difference, or sampling error, occurs by chance, and its variability is measured by

the standard error of the estimate. Sample estimates from a given survey design are unbiased when an average of the estimates from all possible samples would yield, hypothetically, the true population value. In this case, the sample estimate and its standard error can be used to construct approximate confidence intervals, or ranges of values, that include the true population value with known probabilities. If the process of selecting a sample from the population were repeated many times and an estimate and its standard error calculated for each sample then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the true population value.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the true population value.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimates would include the true population value.

Although the estimating methods used in the CPS do not produce unbiased estimates, biases for most estimates are believed to be small enough so that these confidence interval statements are approximately true.

Standard error estimates computed using generalized variance functions are provided in *Employment and Earnings* and other publications. Using replicate variance techniques, standard error estimates are generated. As computed, these standard error estimates reflect contributions not only from sampling error, but from some types of nonsampling error, particularly response variability. Because replicate variance techniques are somewhat cumbersome, simplified formulas called generalized variance functions (GVF's) have been developed for various types of labor force characteristics. The GVF can be used to approximate an estimate's standard error, but this only indicates the general magnitude of its standard error rather than a precise value.

Technical References

Butz, W. P. and Plewes, T. J. (1989) "A Current Population Survey for the 21st Century," *Proceedings of the Fifth Annual Research Conference*, U.S. Department of Commerce, Bureau of the Census: Washington, DC, pp. 3–13.

Bregger, John E. "The Current Population Survey: A Historical View and BLS' Role," *Monthly Labor Review*, June 1984.

Cohany, Sharon R., Polivka, Anne E., and Rothgeb, Jennifer M. "Revisions in the Current Population Survey Effective January 1994," *Employment and Earnings*, February 1994.

Dagum, Estela Bee. *The X–11 ARIMA Seasonal Adjustment Method*. Ottawa, Statistics Canada, January 1983 (Statistics Canada Catalogue No. 12–564E).

National Commission on Employment and Unemployment Statistics. *Counting the Labor Force*, 1979.
A comprehensive review of the entire labor market data system; includes an appraisal of current concepts and methodology and recommendations for further research and improvements.

"Overhauling the Current Population Survey," *Monthly Labor Review*, September 1993.
A trilogy of articles that provides additional information on the background of the CPS redesign.

President's Committee to Appraise Employment and Unemployment Statistics. *Measuring Employment and Unemployment*, 1962.
A review of all Federal statistical series on employment and unemployment and a comparison of the sources and uses of each series; includes a brief history of the development of labor force statistics, an evaluation of concepts and techniques, and recommendations for further research and improvements.

U.S. Department of Commerce, Office of Federal Statistical Policy and Standards. "An Error Profile: Employment as Measured by the Current Population Survey," *Statistical Policy Working Paper* 3, 1978.
A description of the potential sources of error in the CPS as they affect the national employment statistics.

U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings*, Explanatory Notes and Estimates of Error, monthly.
An up-to-date, concise description of the concepts and methods used in the labor force data from the Current Population Survey. Provides tables which present sampling errors for various labor force series.

U.S. Department of Labor, Bureau of Labor Statistics, *Geographic Profile of Employment and Unemployment, 1994*, Bulletin 2469, 1995.
A tabular presentation of 1994 annual average data on the labor force characteristics of the population residing in the Census regions, the 50 States and the District of Columbia, 50 large metropolitan areas, and 17 central cities.

U.S. Department of Labor, Bureau of Labor Statistics, *How the Government Measures Unemployment*, Report 864, 1994.
A short, nontechnical discussion of the concepts and methods used in obtaining labor force statistics from the Current Population Survey.